



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE United States Patent and Trademark Office Address: COMMISSIONER FOR PATENTS P.O. Box 1450 Alexandria, Virginia 22313-1450 www.uspto.gov

APPLICATION NO.	FILING DATE		FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/944,474	08/31/2001		Dan Gallivan	013.0171.01	2388
22895	7590	02/24/2005		EXAMINER	
PATRICK J S INOUYE P S 810 3RD AVENUE				LY, ANH	
SUITE 258	LNOL			ART UNIT	PAPER NUMBER
SEATTLE, WA 98104			2162		
				DATE MAIL ED: 02/24/2004	•

Please find below and/or attached an Office communication concerning this application or proceeding.

Application No. 09/944,474 GALLIVAN ET AL. Examiner Anh Ly The MAILING DATE of this communication appears on the cover sheet with the correspondence address Period for Reply A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION. Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (b) MONTHS from the mailing date of this communication. If the period for reply is specified above, the maximum statutory period will apply and will expire SIX (b) MONTHS from the mailing date of this communication. Failure to reply within the set or extended period for reply vial, by statute, cause the application become ABANDONDE (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b). Status 1) Responsive to communication(s) filed on 27 August 2004. 2a) This action is FINAL. 2b) This action is non-final. 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under Ex parte Quayle, 1935 C.D. 11, 453 O.G. 213.							
## Examiner ## Art Unit ## 2162 The MAILING DATE of this communication appears on the cover sheet with the correspondence address Period for Reply A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION. Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication. If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely. If NO peniod for reply is specified above the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication. Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b). Status 1) ★ Responsive to communication(s) filed on 27 August 2004. 2a) ★ This action is FINAL. 2b) This action is non-final. 3) ★ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under Ex parte Quayle, 1935 C.D. 11, 453 O.G. 213. Disposition of Claims							
The MAILING DATE of this communication appears on the cover sheet with the correspondence address Period for Reply A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b). Status 1) Responsive to communication(s) filed on 27 August 2004. 2a) This action is FINAL. 2b) This action is non-final. 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under Ex parte Quayle, 1935 C.D. 11, 453 O.G. 213.							
The MAILING DATE of this communication appears on the cover sheet with the correspondence address Period for Reply A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION. - Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication. - If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely. - If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication. - Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b). Status 1) Responsive to communication(s) filed on 27 August 2004. 2a) This action is FINAL. 2b) This action is non-final. 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under Ex parte Quayle, 1935 C.D. 11, 453 O.G. 213.							
A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION. - Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication. - If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely. - If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication. - Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b). Status 1) ★ Responsive to communication(s) filed on 27 August 2004. 2a) ★ This action is FINAL. 2b							
after SIX (6) MONTHS from the mailing date of this communication. If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely. If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication. Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b). Status 1) Responsive to communication(s) filed on 27 August 2004. 2a) This action is FINAL. 2b) This action is non-final. 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under Ex parte Quayle, 1935 C.D. 11, 453 O.G. 213. Disposition of Claims							
1)⊠ Responsive to communication(s) filed on <u>27 August 2004</u> . 2a)⊠ This action is FINAL . 2b)□ This action is non-final. 3)□ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under <i>Ex parte Quayle</i> , 1935 C.D. 11, 453 O.G. 213. Disposition of Claims							
2a) ☐ This action is FINAL. 2b) ☐ This action is non-final. 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under Ex parte Quayle, 1935 C.D. 11, 453 O.G. 213. Disposition of Claims							
3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under <i>Ex parte Quayle</i> , 1935 C.D. 11, 453 O.G. 213. Disposition of Claims							
closed in accordance with the practice under <i>Ex parte Quayle</i> , 1935 C.D. 11, 453 O.G. 213. Disposition of Claims							
4)⊠ Claim(s) <u>1-44</u> is/are pending in the application.							
4a) Of the above claim(s) is/are withdrawn from consideration.							
5) Claim(s) is/are allowed.							
6)⊠ Claim(s) <u>1-44</u> is/are rejected.							
7) Claim(s) is/are objected to.							
8) Claim(s) are subject to restriction and/or election requirement.							
Application Papers							
9) The specification is objected to by the Examiner.							
10)⊠ The drawing(s) filed on <u>31 August 2001</u> is/are: a)⊠ accepted or b)□ objected to by the Examiner.							
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).							
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR-1.121(d).							
11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.							
Priority under 35 U.S.C. § 119							
12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f). a) All b) Some * c) None of:							
1. Certified copies of the priority documents have been received.							
2. Certified copies of the priority documents have been received in Application No							
3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).							
* See the attached detailed Office action for a list of the certified copies not received.							
Attachment(s)							
1) Notice of References Cited (PTO-892) 4) Interview Summary (PTO-413) 2) Notice of Draftsperson's Patent Drawing Review (PTO-948) Paper No(s)/Mail Date							
2) Notice of Draftsperson's Patent Drawing Review (PTO-948) 3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date 5) Notice of Informal Patent Application (PTO-152) 6) Other:							

DETAILED ACTION

1. This Office Action is response to Applicants' response filed on 08/27/2004.

Response to Arguments

- 2. Applicant's arguments filed on 08/27/2004 have been fully considered but they are not persuasive.
- Applicants argued that, "A prima facie case of obviousness has not been shown."
 (Page 12, 3rd paragraph, Page 15, lines 12-25).

In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, Marchisio (US PATNO.: 6,510,406) and Holt et al. (US PATNO.: 6,701,305, hereinafter Holt) are from the same field of endeavor and both are directed to analyzing of the collection of documents for conceptual relationships, concepts of documents and term and text

Page 3

Art Unit: 2162

mining. One having ordinary skill in the art would have found it motivated to combine the teachings of Marchisio and Holt because that would provide Marchisio's system enhanced capability of retrieving information that deals with the problem of synonymy, polysemy and retrieval by concept, thereby generating and determining a frequency of occurrences of concepts in the collection of documents. Moreover, the examiner kindly submits that the applicants misread the applicable references used in the last office action. However, when read and analyzed in light the specification, the invention as claimed does not support applicant's assertions. Actually, applicants are interpreting the claims very narrow without considering the broad teaching of the references used in the rejections. Additionally, it is important to note that the examiner interpretation of the claims, wherein, the examiner explicitly stated passages in the cited references which were not even addressed. The aforementioned assertion wherein all the limitations are not taught or suggested by the prior art of the record, was unsupported by objective factual evidence and was not found to be substantial evidentiary value. The examiner has provided in the last office action, a convincing one of reasoning as to why the artisan would have found the claimed invention to have been obvious in light of the teachings of the cited references. Applicants are reminded that 37 CFR 1.111(b) states, a general allegation that the claims define a patentable invention without specifically printing out how the language of the claims patentably distinguishes them from the references does not comply with the requirements of this section. Therefore, the applicants have failed to provided prima facie evidence how the language of the claims

patentably distinguished them from the cited references. Hence, the applicants' assertions are just mere allegation with no supported fact.

Applicants argued that, "In combination, the teachings of Marchisio with the teachings of Holt, does not teach or suggest all limitations of the claimed invention." (Page 14, lines 1-2, 12-14, 28-30, Page 16, lines 25-26, Page 15, lines 10-12, Page 17, lines 10-12, 18-20, 24-25 and Page 18, lines 4-6). Marchisio teaches an information retrieval system that deals with the problem of synonymy, polysemy and retrieval of concept, determining a frequency of occurrences of concepts of the phrase of an electronic document, the proximity of such concepts (col. 7, lines 50 60), cluster data and establish conceptual links, indexing the document by transposing individual document into a term-document matrix (col. 9, lines 10-15 and lines 45-67). Also, Marchisio teaches determines the dimensionality of the transform space, the number of concepts in transform space, and the coarser the cluster of documents and that relevance weights are assigned more uniformly across a document collection (col. 15, lines 1-16). While Holt teaches distance calculation based on the document set, compute term-by-document matrix and to weight the terms (term weighting) in order to reduce the impact of high frequency terms (see figs. 3 and col. 5, lines 25-35 and fig. 8), and the relationship of concepts in the set of documents in order to obtain more meaningful results (col. 8, lines 40-48) and the plurality of documents represented by the subspace representation can be scored with respect to the query and the documents can then be ranked in terms of score with the more relevant documents

Application/Control Number: 09/944,474 Page 5

Art Unit: 2162

having a better score. As a result of the weighting of those portions of the subspace representation that relate to the terms of the query, the resulting scores can be more meaningful (col. 14, lines 28-50).

Thus, Applicants' arguments are not persuasive over the prior art of the record.

Claim Rejections - 35 USC § 103

- 4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:
 - (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.
- 5. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

6. Claims 1-44 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent No. 6,510,406 issued to Marchisio in view of US Patent No. 6,701, 305 issued to Holt et al. (hereinafter Holt).

With respect to claim 1, Marchisio teaches a histogram module determining a frequency of occurrences of concepts in a set of unstructured documents, each concept representing an element occurring in one or more of the unstructured documents (the frequency of occurrences of individual terms based on the extraction based on concepts: see fig. 1 and col. 45-65);

a selection module selecting, a subset of concepts out of the frequency of occurrences (parsing the user query into terms or phrases and the proximity based on the concepts: col. 7, lines 50-58),

grouping one or more concepts from the concepts subset (grouping the concepts: col. 15, lines 60-67 and col. 16, lines 1-8; also see fig. 7):

assigning weights to one or more clusters of concepts for each group of concepts (assigning weight to the term in the user query : col. 15, lines 1-16 and col. 2, lines 25-42); and

each document indexed by each such group of concepts between the frequency of occurrences and the weighted cluster (col. 15, lines 1-16 and col. 9, lines 52-67 and col. 10, lines 1-30).

Marchisio discloses searching or retrieving by latent concept or latent semantic for the fundamental problems of synonymy and polysemy in the text mining and using data mining techniques in order to overcome a wide margin of uncertainty in the initial

choice of a keyword in a query, from which the user can query unstructured document such as electronic message or document) or structured document such as document storing in the database with indexing, clustering of documents with the concepts (col. 15, lines 1-16, see abstract, fig. 2) and indexing document (col. 9, lines 52-67 and col. 10, lines 1-30). Marchisio also teaches the computational approximation of query (fig. 6 and col. 15, lines 25-50). Marchisio does not explicitly teach a best-fit module calculating a best-fit approximation for each document for each such concept grouped into the group of concepts.

However Holt teaches approximating some of the semantics latent in the documents for the synonymy and polysemy in the documents (col. 3, lines 40-67 and col. 4, lines 1-40).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Marchisio with the teachings of Holt, wherein the computing the similarity between each of the electronic information files (Marchisio's fig. 1 and col. 8, lines 8-18), would incorporate the user of calculating approximation for each document, in the same conventional manner as disclosed by Holt (col. 3, lines 40-67 and col. 4, lines 1-40). The motivation being to provide the system having a determination for conceptual relationships in the text mining and for easing to deal with the problems of synonymy and polysemy.

With respect to claims 2-3, Marchisio teaches an extraction module extracting features from each of the unstructured documents and normalizing the extracted features into the concepts and a structured database storing the extracted features as

uniquely identified records (see fig. 2, feature extraction from a relational database and col. 9, lines 8, lines 33-67 and col. 9, lines 1-15).

With respect to claim 4, Marchisio teaches a visualization module visualizing the frequency of occurrences, comprising at least one of creating a histogram mapping the frequency of occurrences for each document in the unstructured documents set and creating a corpus graph mapping the frequency of occurrence for all such documents in the unstructured documents set (see figs 8-9. col. 16, lines 33-58).

With respect to claim 5, Marchisio teaches a threshold comprising a median and edge conditions, each such concept in the concepts subset occurring within the edge conditions (col. 6, lines 55-65 and col. 7, lines 18-27 and fig. 1).

With respect to claims 6-7, Marchisio teaches an inner product module determining, for each group of concepts, the best fit approximation as the inner product between the frequency of occurrences and the weighted cluster for each such concept in the group of concepts, and wherein the inner product d (cluster) is calculated according to the equation comprising: d(cluster) = ~ doc(termi) * cluster (termi) where doc(concept) represents the frequency of occurrence for a given concept in the document and cluster(concept) represents the weight for a given cluster (inner product's computation and its equation: col. 2, lines 3-42 and approximation for the query: fig. 6 and col. 15, lines 25-50).

With respect to claim 8, Marchisio teaches a system as discussed in claim 1.

Marchisio discloses searching or retrieving by latent concept or latent semantic for the fundamental problems of synonymy and polysemy in the text mining and using

data mining techniques in order to overcome a wide margin of uncertainty in the initial choice of a keyword in a query, from which the user can query unstructured document such as electronic message or document) or structured document such as document storing in the database with indexing, clustering of documents with the concepts (col. 15, lines 1-16, see abstract, fig. 2) and indexing document (col. 9, lines 52-67 and col. 10, lines 1-30). Marchisio also teaches the computational approximation of query (fig. 6 and col. 15, lines 25-50). Marchisio does not explicitly teach a control module iteratively re-determining the best-fit approximation to a change in the set of unstructured documents.

However Holt teaches approximating some of the semantics latent in the documents for the synonymy and polysemy in the documents (col. 3, lines 40-67 and col. 4, lines 1-40).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Marchisio with the teachings of Holt, wherein the computing the similarity between each of the electronic information files (Marchisio's fig. 1 and col. 8, lines 8-18), would incorporate the user of calculating approximation for each document, in the same conventional manner as disclosed by Holt (col. 3, lines 40-67 and col. 4, lines 1-40). The motivation being to provide the system having a determination for conceptual relationships in the text mining and for easing to deal with the problems of synonymy and polysemy.

Art Unit: 2162

Claim 9 is essentially the same as claim 1 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 1 hereinabove.

Claim 10 is essentially the same as claim 2 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 2 hereinabove.

Claim 11 is essentially the same as claim 3 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 3 hereinabove.

Claim 12 is essentially the same as claim 4 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 4 hereinabove.

Claim 13 is essentially the same as claim 5 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 5 hereinabove.

Claim 14 is essentially the same as claim 6 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 6 hereinabove.

Claim 15 is essentially the same as claim 7 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 7 hereinabove.

Claim 16 is essentially the same as claim 8 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 8 hereinabove.

Claim 17 is essentially the same as claims, 9, 10, 11, 12, 13, 14, 15 or 16 except that it is directed to a computer-readable storage medium rather than a method, and is rejected for the same reason as applied to the claims 9, 10, 11, 12, 13, 14, 15 or 16 hereinabove.

With respect to claim 18, Marchisio teaches an extraction module extracting a multiplicity of concepts from a set of unstructured documents into a lexicon uniquely identifying each concept and a frequency of occurrence (see fig. 2, item 21 and 28, and also see fig. 1, item 16, col. 7, lines 28-38": parsing the unstructured text into lexical uniquely identifier for each concept (concept ID number); the frequency of occurrences of individual terms based on the extraction based on concepts: see fig. 1 and col. 45-65);

a frequency-mapping module creating a frequency of occurrence representation for each documents set, the representation providing an ordered corpus of the frequencies of occurrence of each concept (best match model: col. 2, lines 3-56 and col. 8, lines 20-32);

a concept selection module selecting a subset of concepts from the frequency of occurrence representation filtered against a minimal set of concepts each referenced in at least two documents with no document in the corpus being unreferenced (parsing the user query into terms or phrases and the proximity based on the concepts: col. 7, lines

50-58; a set of documents to be inspected and some not: to be searched or extracted: col. 11, lines 20-46);

a group generation module generating a group of weighted clusters of concepts selected from the concepts subset (grouping the concepts: col. 15, lines 60-67 and col. 16, lines 1-8; also see fig. 7); and

each document weighted against each group of weighted clusters of concepts (col. 15, lines 1-16 and col. 9, lines 52-67 and col. 10, lines 1-30)

Marchisio discloses searching or retrieving by latent concept or latent semantic for the fundamental problems of synonymy and polysemy in the text mining and using data mining techniques in order to overcome a wide margin of uncertainty in the initial choice of a keyword in a query, from which the user can query unstructured document such as electronic message or document) or structured document such as document storing in the database with indexing, clustering of documents with the concepts (col. 15, lines 1-16, see abstract, fig. 2) and indexing document (col. 9, lines 52-67 and col. 10, lines 1-30). Marchisio also teaches the computational approximation of query (fig. 6 and col. 15, lines 25-50). Marchisio does not explicitly teach a best-fit module calculating a best-fit approximation for each document for each such concept grouped into the group of concepts.

However Holt teaches approximating some of the semantics latent in the documents for the synonymy and polysemy in the documents (col. 3, lines 40-67 and col. 4, lines 1-40).

Art Unit: 2162

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Marchisio with the teachings of Holt, wherein the computing the similarity between each of the electronic information files (Marchisio's fig. 1 and col. 8, lines 8-18), would incorporate the user of calculating approximation for each document, in the same conventional manner as disclosed by Holt (col. 3, lines 40-67 and col. 4, lines 1-40). The motivation being to provide the system having a determination for conceptual relationships in the text mining and for easing to deal with the problems of synonymy and polysemy.

With respect to claims 19-22, Marchisio teaches a histogram module creating a histogram mapping the frequency of occurrence representation for each document in the documents set (generating the term-document matrix to indicating the number of occurrences of the term: see fig. 1, item 6, col. 6, lines 35-65; the frequency of occurrences of individual terms based on the extraction based on concepts: see fig. 1 and col. 45-65);

a data mining module mining the multiplicity of concepts from each document as at least one of a noun, noun phrase and tri-gram (col. 2, lines 57-67);

a normalizing module normalizing the multiplicity of concepts into a substantially uniform lexicon (10, 10-31); and

wherein the substantially uniform lexicon is in third normal form (col. 12, lines 1-36).

Art Unit: 2162

With respect to claim 23, Marchisio teaches a corpus mapping module creating a corpus graph mapping the frequency of occurrence representation for all documents in the documents set (see figs 8-9. col. 16, lines 33-58).

With respect to claim 24, Marchisio teaches a threshold module defining the pre-defined threshold as a median value and a set of edge conditions and choosing those concepts falling within the edge conditions as the concepts subset (col. 6, lines 55-65 and col. 7, lines 18-27 and fig. 1).

With respect to claim 25-26, Marchisio teaches a cluster module naming, one or more of the concepts within the concepts subset to a cluster and assigning a weight to each concept with each such cluster (assigning weight to the term in the user query: col. 15, lines 1-16 and col. 2, lines 25-42; filtering and identifying the cluster: col. 3, lines 5-12); and

.a group module grouping, one or more of the clusters into each such group of weighted clusters of concepts (assigning weight to the term in the user query : col. 15, lines 1-16 and col. 2, lines 25-42 and partitioning or grouping of documents: col. 15, lines 40-50).

With respect to claim 27, Marchisio teaches a system as discussed in claim 18. Marchisio discloses searching or retrieving by latent concept or latent semantic for the fundamental problems of synonymy and polysemy in the text mining and using data mining techniques in order to overcome a wide margin of uncertainty in the initial choice of a keyword in a query, from which the user can query unstructured document such as electronic message or document) or structured document such as document storing in

Art Unit: 2162

the database with indexing, clustering of documents with the concepts (col. 15, lines 1-16, see abstract, fig. 2) and indexing document (col. 9, lines 52-67 and col. 10, lines 1-30). Marchisio also teaches the computational approximation of query (fig. 6 and col. 15, lines 25-50) and building the computation of the distance between the query vector and document clusters in the optimization problem. Marchisio does not explicitly teach a Euclidean module calculating a Euclidean distance between the frequency of occurrence for each document and a corresponding weighted cluster.

However Holt teaches the Euclidean distance of the vector (col. 4, lines 25-48).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Marchisio with the teachings of Holt, wherein the computing the similarity between each of the electronic information files (Marchisio's fig. 1 and col. 8, lines 8-18), would incorporate the user of calculating approximation for each document, in the same conventional manner as disclosed by Holt (col. 3, lines 40-67 and col. 4, lines 1-40). The motivation being to provide the system having a determination for conceptual relationships in the text mining and for easing to deal with the problems of synonymy and polysemy.

With respect to claim 28, Marchisio teaches a iteration module removing select documents from the documents set and iteratively reevaluating the matrix of best fit approximations based on a revised frequency of occurrence representation and concepts subset (col. 14, lines 56-67; and removing the documents; col. 7, lines 55-65 and col.4, lines 30-36).

Art Unit: 2162

With respect to claims 29-30, Marchisio teaches a structured database storing the lexicon, the lexicon comprising a plurality of records each uniquely identifying one such concept and an associated frequency of occurrence (see fig. 1 and fig 2, col. 9, lines 8-14); and

wherein the structured database is an SOL database (col. 10, lines 32-45 and lines 58-67 and col. 11, lines 1-5, see fig. 2).

Claim 31 is essentially the same as claim 18 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 18 hereinabove.

Claim 32 is essentially the same as claim 19 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 19 hereinabove.

Claim 33 is essentially the same as claim 20 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 20 hereinabove.

Claim 34 is essentially the same as claim 21 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 21 hereinabove.

Claim 35 is essentially the same as claim 22 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 22 hereinabove.

Art Unit: 2162

Claim 36 is essentially the same as claim 23 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 23 hereinabove.

Claim 37 is essentially the same as claim 24 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 24 hereinabove.

Claim 38 is essentially the same as claim 18 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 25 hereinabove.

Claim 39 is essentially the same as claim 18 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 26 hereinabove.

Claim 40 is essentially the same as claim 18 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 27 hereinabove.

Claim 41 is essentially the same as claim 18 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 28 hereinabove.

Claim 42 is essentially the same as claim 18 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 29 hereinabove.

Page 18

Application/Control Number: 09/944,474

Art Unit: 2162

Claim 43 is essentially the same as claim 18 except that it is directed to a method rather than a system, and is rejected for the same reason as applied to the claim 30 hereinabove.

Claim 44 is essentially the same as claims, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42 except that it is directed to a computer-readable storage medium rather than a method, and is rejected for the same reason as applied to the claims, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42 hereinabove.

Application/Control Number: 09/944,474 Page 19

Art Unit: 2162

Conclusion

7. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Page 20

Application/Control Number: 09/944,474

Art Unit: 2162

Contact Information

8. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Anh Ly whose telephone number is (571) 272-4039 or via E-Mail: ANH.LY@USPTO.GOV or fax to (571) 273-4039. The examiner can normally be reached on TUESDAY – THURSDAY from 8:30 AM – 3:30 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene, can be reached on (571) 272-4107 or Primary Examiner Jean Corrielus (571) 272-4032.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Any response to this action should be mailed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, or faxed to: Central Fax Center (703) 872-9306

JEAN M. CORRIELUS PRIMARY EXAMINER